

(43) Date of A Publication 17.02.1999

(21) Application No **9814880.2**

(22) Date of Filing **10.07.1998**

(30) Priority Data

(31) **08893470** (32) **11.07.1997** (33) **US**

(71) Applicant(s)

**International Business Machines Corporation  
(Incorporated in USA - New York)  
Armonk, New York 10504, United States of America**

(72) Inventor(s)

**Joe C Dubbels  
Kevin P Gibson**

(74) Agent and/or Address for Service

**S R Davies  
IBM United Kingdom Limited, Intellectual Property  
Dept, Hursley Park, Winchester, Hampshire,  
SO21 2JN, United Kingdom**

(51) INT CL<sup>6</sup>

**G06F 17/30**

(52) UK CL (Edition Q )

**G4A AUDB**

(56) Documents Cited

**GB 2305525 A EP 0851367 A1  
Seybold Report on Internet Publishing Vol. 2 No. 1,  
September 1997, pages 5-20 & IAC Acc No 19756197**

(58) Field of Search

**UK CL (Edition P ) G4A AUDB  
INT CL<sup>6</sup> G06F 17/30  
Online: WPI, EDOC, INSPEC, COMPUTER**

(54) Abstract Title

**Printing pages related to a viewed WWW page**

(57) Enabling a user to select a web page, then print some or all of the linked pages which, based on predetermined criteria, are related to that page, without having first to invoke the linked pages. Each web page includes an applet (310, fig. 3) that runs on the web client system (200, fig. 3) when a print button on the page is pressed, 420. A print tool (330, fig. 3) running on the server (220, fig. 3) then parses the selected page and builds a list of related pages, 430, and allows the user to select which of the related pages will be printed, 440, 450. The print tool then constructs a temporary web page containing all the selected web pages, 460. This temporary web page is printed, 480, using the browser standard print function (320, fig. 3). Alternatively, a print utility in the web client performs the parsing and list building functions (figs. 5 and 6).

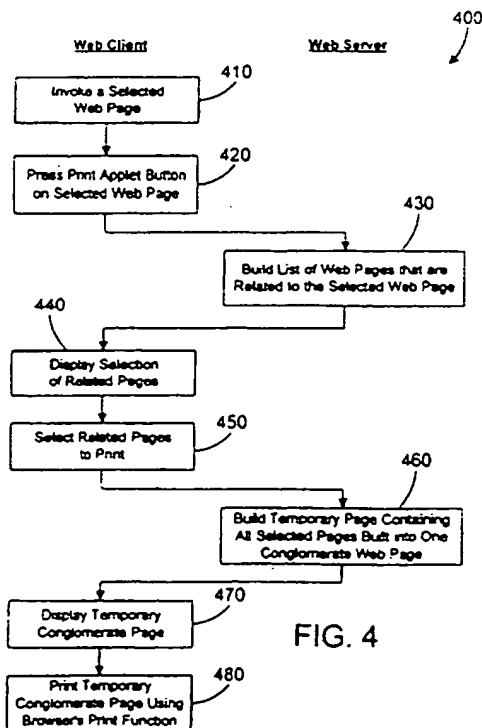


FIG. 4

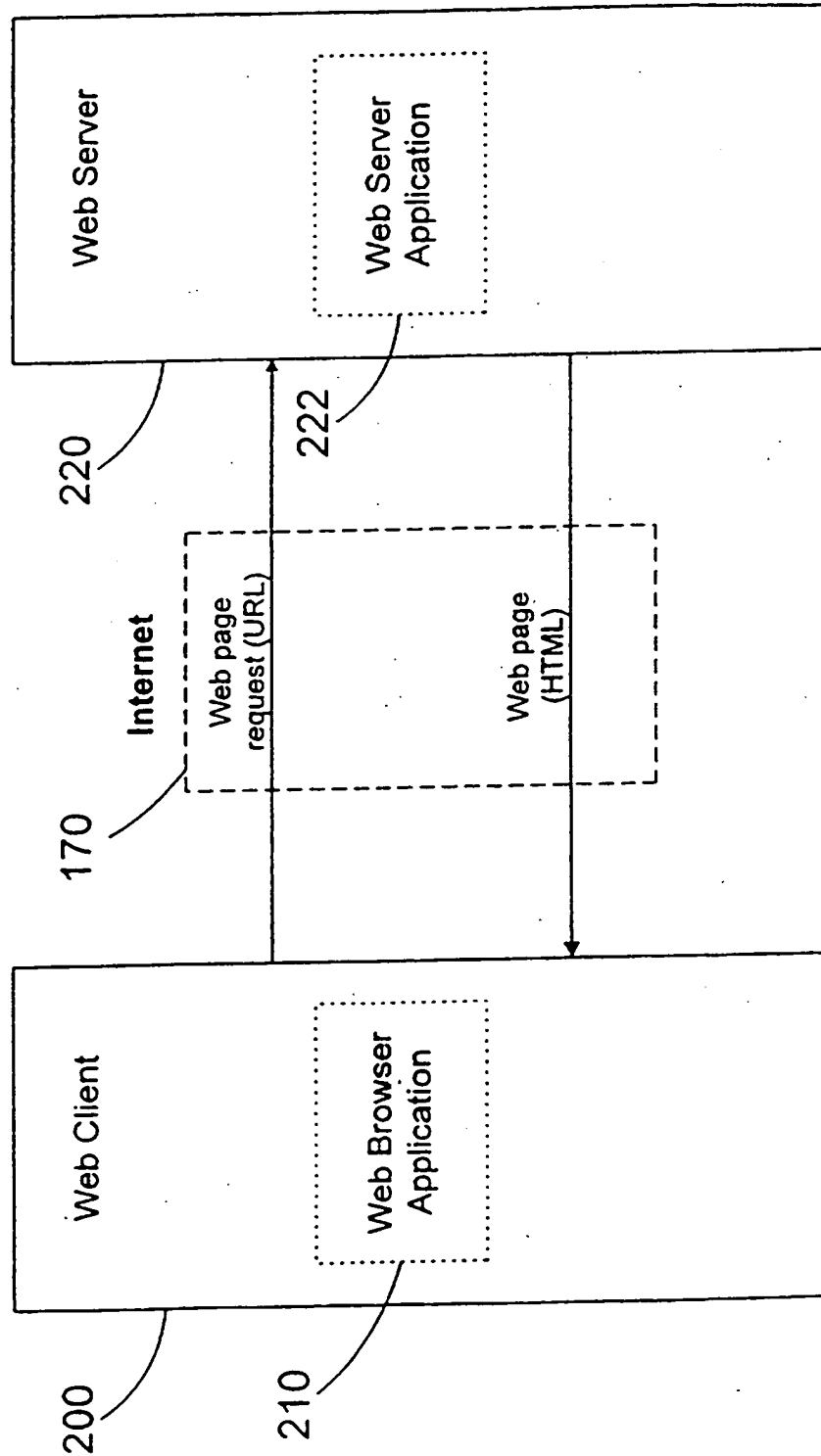
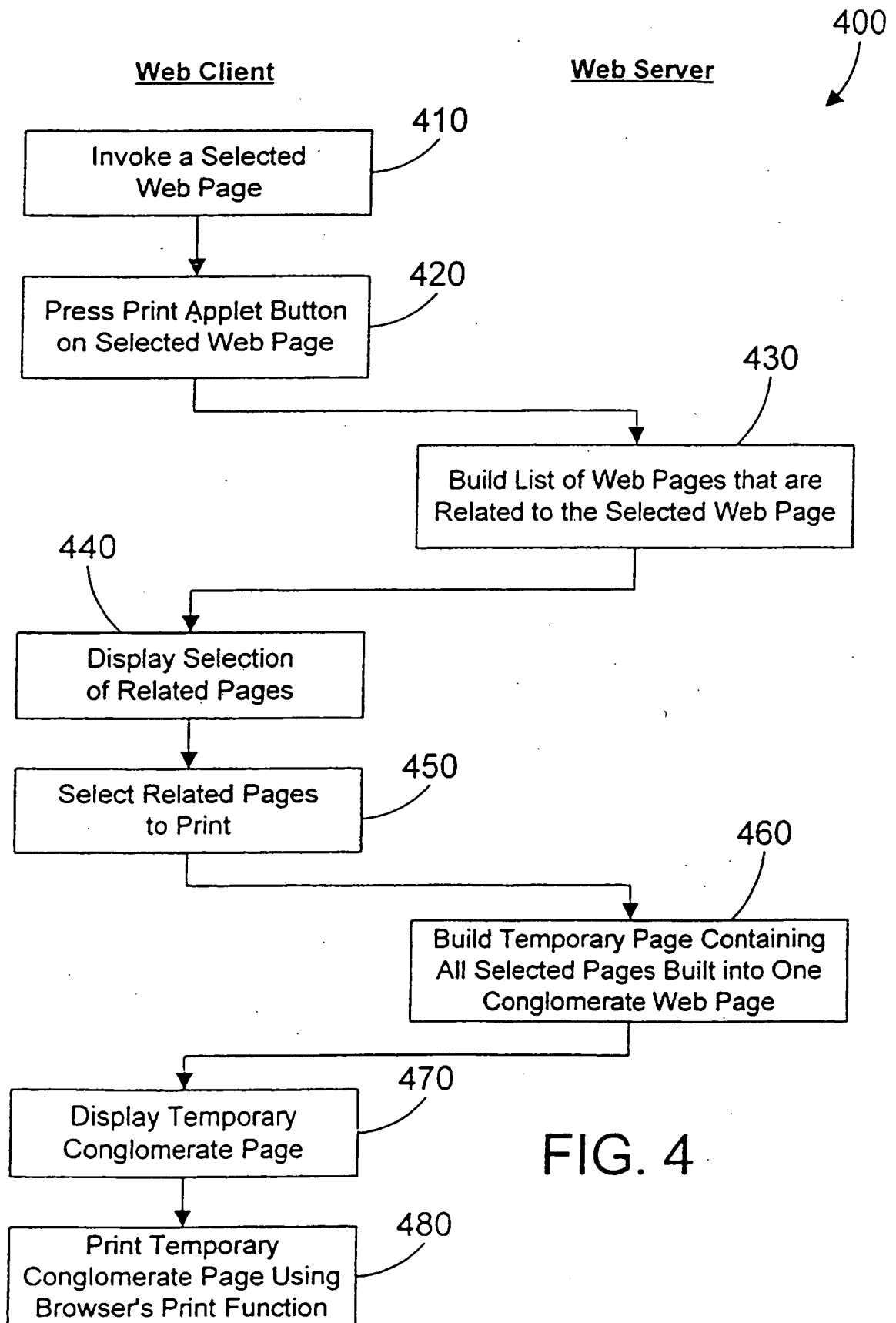


FIG. 2



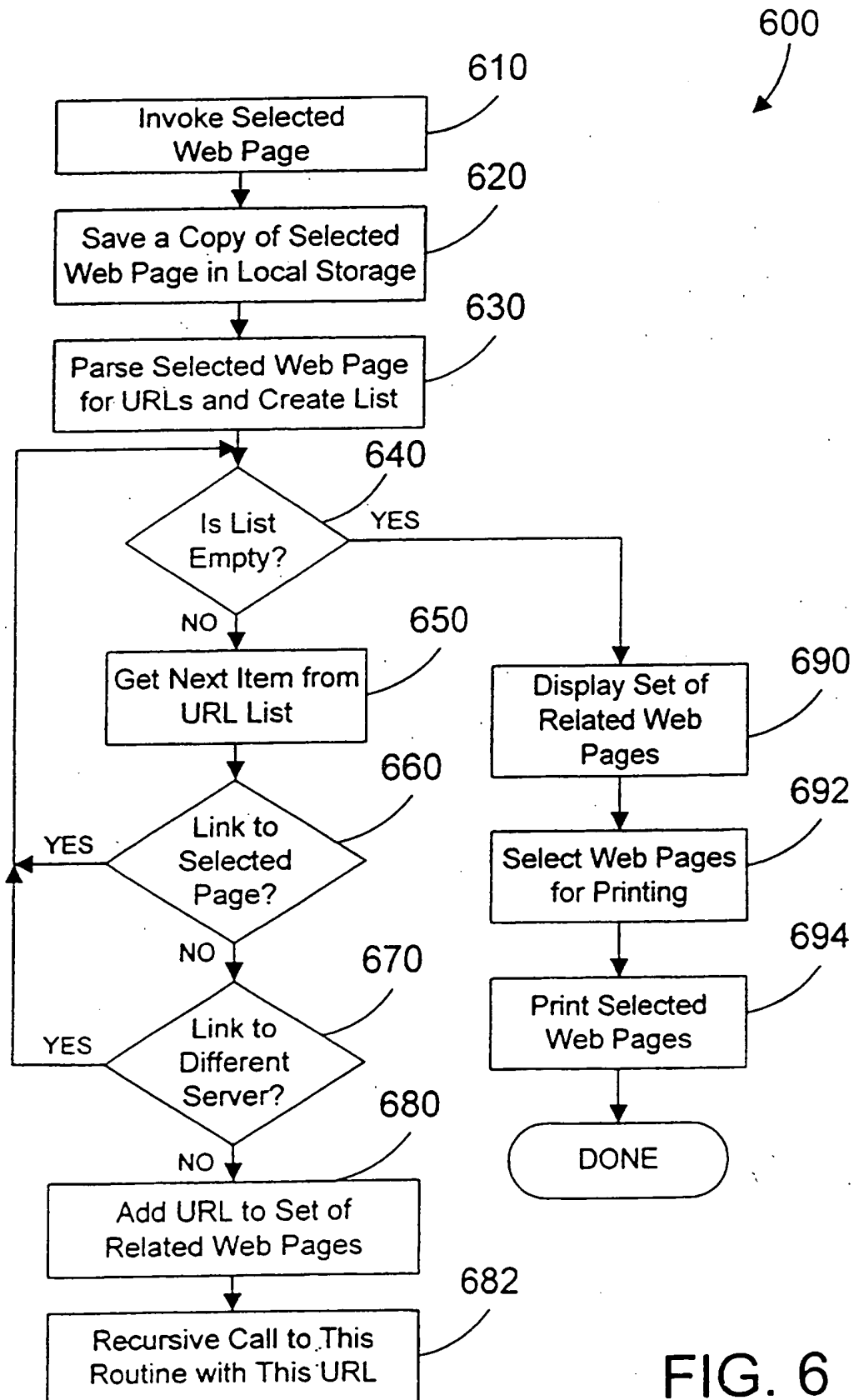


FIG. 6

800

810		
	Web Page URL	Web Page Title
✓	www.walt.com	Walt's World of Melons
	www.walt.com\company	Walt's Company Information
✓	www.walt.com\pricing	Walt's Pricing and Ordering
✓	www.walt.com\delivery	Walt's Delivery Terms and Conditions

820

Print Checked Pages

FIG. 8

and print each of the 40 web pages. Needless to say, this process becomes very time-consuming. As the number of Internet users, providers, and web servers continues to rapidly expand, such problems in the printing of web pages will continue to be an impediment to the effective usage of resources available on the Internet.

Accordingly, the invention provides apparatus for printing multiple pages comprising:

means for selecting a page;

means for automatically determining one or more pages related to said selected page;

and means for printing said selected page and said one or more related pages.

Thus the convenience of printing related web pages or other suitable material is improved by providing ways that a user may print related pages without the customary user interaction required to invoke and print each page.

In a preferred embodiment, the apparatus further comprises a page parsing and listing mechanism, and the selected page is in hypertext markup language (HTML) and is selected using a Uniform Resource Locator (URL). A page can be considered as related to said selected page if they both reside on the same server, and optionally if they also both have the same base address, or using any other suitable criteria.

In one embodiment, said apparatus comprises a computer program product comprising computer program instructions recorded on a storage medium. In an alternative embodiment, said apparatus comprises a computer system including at least one processor, a memory coupled to the at least one processor; and computer program instructions in said memory for execution by the program to perform said printing operation. Moreover, it will be appreciated that the apparatus may comprise any suitable configuration of hardware and software, whether the apparatus is a single computer system or is comprised of multiple computer systems operating in concert.

The invention further provides apparatus including a web page print mechanism comprising:

a web page selection mechanism that allows a user to select at least one web page from the list of web pages;

a web page print mechanism residing in the memory and executed by the at least one processor, the web page print mechanism comprising:

a web page parsing and listing mechanism that generates a list of web pages related to the selected web page;

a web page selection mechanism that allows a user to select at least one web page from the list of web pages; and

a mechanism for printing the at least one web page selected by the user using the web page selection mechanism.

Viewed from another aspect the invention further provides a program product comprising:

(A) a print mechanism, the print mechanism printing a plurality of pages that are related to a selected page;

(B) signal bearing media bearing the print mechanism.

Such a program product may be distributed in a variety of forms, without detriment to operation in accordance with the present invention, regardless of the particular type of signal bearing media used to actually carry out the distribution. Examples of suitable signal bearing media include: recordable type media such as floppy disks and CD ROM, and transmission type media such as digital and analog communications links.

In a preferred embodiment, the print mechanism comprises a web page print mechanism including:

a web page parsing and listing mechanism that generates a list of web pages related to a selected web page; and

a web page selection mechanism that allows a user to select from the list of web pages at least one web page to be printed.

Thus the apparatus and method described herein for printing related web pages allow a web user to select a web page, then print all of the related web pages based on one or more predetermined criteria. A web user is therefore able to print related web pages without manually invoking and printing each page. In a first embodiment, each web page includes an applet that is run on the web client system when a print button on the page is pressed. The client applet communicates with a print tool running on the server that parses the selected page and builds a list of related pages and allows the user to select which of the related pages will be printed. Once the user selects the pages to be printed, the print tool constructs a temporary web page that contains all the web pages the user selected. This temporary web page may then be

A web page is primarily visual data that is intended to be displayed on the monitor of user workstation 200. Web pages are generally written in Hypertext Markup Language (HTML). When web server 220 receives a web page request, it will build a web page in HTML and send it off across the Internet 170 to the requesting web browser 210. Web browser 210 understands HTML and interprets it and outputs the web page to the monitor of user workstation 200. This web page displayed on the user's screen may contain text, graphics, and links (which reference addresses of other web pages.) These other web pages (i.e., those represented by links) may be on the same or on different web servers. The user can go to these other web pages by clicking on these links using a mouse or other pointing device. This entire system of web pages with links to other web pages on other servers across the world is known as the "World Wide Web".

Referring now to FIG. 1, a computer system 100 includes a processor 110, a main memory 120, a mass storage interface 140, and a network interface 150, all connected by a system bus 160. Those skilled in the art will appreciate that this system encompasses all types of computer systems: personal computers, midrange computers, mainframes, etc. Note that many additions, modifications, and deletions can be made to this computer system 100; examples of possible additions include: a computer monitor, a keyboard, a cache memory, and peripheral devices such as printers.

Processor 110 can be constructed from one or more microprocessors and/or integrated circuits. Processor 110 executes program instructions stored in main memory 120. Main memory 120 stores programs and data that the computer may access. When computer system 100 starts up, processor 110 initially executes the program instructions that make up operating system 126. Operating system 126 is a sophisticated program that manages the resources of the computer system 100. Some of these resources are the processor 110, main memory 120, mass storage interface 140, network interface 150, and system bus 160.

Main memory 120 includes one or more application programs 122, data 124, operating system 126, a web page print mechanism 128, and one or more web pages 130. Application programs 122 are executed by processor 110 under the control of operating system 126. Application programs 122 can be run with program data 124 as input. Application programs 122 can also output their results as program data 124 in main memory. As



System bus 160 allows data to be transferred among the various components of computer system 100. Although computer system 100 is shown to contain only a single main processor and a single system bus, those skilled in the art will appreciate that embodiments may be formed using a computer system that has multiple processors and/or multiple buses. In addition, the interfaces that are used in the preferred embodiment may include separate, fully programmed microprocessors that are used to off-load compute-intensive processing from processor 110, or may include I/O adapters to perform similar functions.

Two different embodiments for printing related web pages will now be described. The first embodiment uses an applet on the web client in conjunction with a print tool that resides on the web server. An example of a suitable system and method in accordance with the first embodiment is shown in FIGS. 3 and 4. Any of the programs executing on a web server are referred to generically herein as web server programs, and any of the programs executing on the web client are referred to generically herein as web client programs. The second embodiment of the present invention does not require any software to be installed on the web server. An example of a suitable system in accordance with the second embodiment is shown in FIGS. 5 and 6.

An apparatus 300 in accordance with the first embodiment is illustrated in FIG. 3, and includes a web client 200 coupled to a web server 220 via the Internet 170. Web client 200 includes a web browser application 210 and a print applet 310. The function of web browser application 210 is described above and is well-known in the art, and includes a web client print mechanism 320 that is used to print individual web pages. Print applet 310 is a small application such as a Java applet that is invoked when a user takes a particular action with respect to a selected web page. In the preferred embodiment, print applet 310 is executed when a user selects a particular "print button" on a web page that includes print applet 310. Print applet 310 is shown to reside on web client 200, but those skilled in the art will recognize that applets such as print applet 310 are typically dynamically loaded from web server 220 to web client 200 with a web page. A user presses the "print button" on a web page that corresponds to print applet 310 to indicate that printing of the current page and its related pages is desired.

pages that share the base address `www.companyX.com/support` are related to the selected page, while other pages at this site are not related. Regardless of the specific criteria used, pages that are related to the selected web page are included in the list, and pages that are not related are not included in the list.

The list of related web pages is then passed to the web client, which displays the list to the user (step 440). The user then selects the pages on the list to print (step 450). The list of selected pages is then passed to the web server, which uses this information to build a temporary web page that is a conglomerate of all the pages that were selected for printing (step 460). The temporary conglomerate web page is built by the web page merging mechanism 350, which performs the necessary functions to convert several individual web pages into a single web page. For example, a tag `<body>` generally defines the beginning of an HTML page, and the tag `</body>` defines the end of an HTML page. For the case of printing HTML pages, web page merging mechanism 350 builds the conglomerate web page by removing the `</body>` tag in the first page to be printed, by removing the `<body>` tag in the last page to be printed, and by removing all `<body>` and `</body>` tags for all pages in between. In addition, other tags such as header and end tags may be moved to the beginning or end of the conglomerate web page, or may be deleted, if appropriate. This results in a single conglomerate web page that contains all the pages to be printed. This conglomerate web page is then passed to the web client and displayed to the user (step 470). The conglomerate web page may then be printed using the conventional print function that is supplied with the web browser application (step 480).

By providing a print applet that is downloaded to a web browser with a web page along with a print tool program running on the web server, a user may print multiple related pages with a standard web browser. This approach requires new software to be added to the web server. In the second embodiment, discussed in more detail below, no additional software is added to the web server. Instead, software is added to the web client to provide the capability of printing multiple related web pages.

Referring now to FIGS. 5 and 6, an apparatus 500 in accordance with the second embodiment includes a web client 200 and a web server 220 connected via the Internet 170. Web client 200 includes a web browser application 210 and a web page print mechanism 128. The web browser

traverses all related pages to build a complete set of pages to be printed.

5       The skilled person will recognise many possible modifications to the approach described above. For example, a threshold may be selected by the user to determine the number of levels deep from the initial selected page method 600 will go in selecting related web pages. A threshold of two, for example, would limit the related pages to those that are one or two links away from the initial selected page. Other  
10       thresholds and criteria for determining what pages are related may also be employed.

15       Once all lists are empty for all URLs considered (step 640=YES), the list of related web pages is displayed to the web user (step 690). The web user then selects which of the related pages are to be printed (step 692), and these selected pages are then printed (step 694). Note that the print function is supplied by the selected web page print mechanism 560 (FIG. 5).

20       The function of the two embodiments disclosed herein may be best understood from a web user's viewpoint as shown in FIGS. 7 and 8. A sample web page 700 is shown that includes an image of a watermelon 710 and an image of a honeydew 730. Each of these two images is represented by a graphical interface file (GIF). Another image of a cantaloupe 720 is  
25       included in another graphical format known as JPG. In addition, two links 740 and 750 are provided to allow a user to go to these related web pages. An additional button 760 is provided, which corresponds to the button used to run print applet 310 in the preferred embodiment.

30       If web page 700 is the first page invoked by method 400 or method 600, the links in the page will be parsed and put in a list of related links. In this specific example, the two links 740 and 750 are related links, but any number of unrelated links may also be added on page 700. Unrelated links are ignored, so only related links are considered for  
35       this particular example.

40       We assume for the specific example of FIGS. 7 and 8 that the web page Walt's Company Information corresponding to link 740 has no links and that the web page Walt's Pricing and Ordering corresponding to link 750 has one link to another related page entitled Walt's Delivery Terms and Conditions. Thus, when all of the related links are parsed and a

CLAIMS

1. Apparatus for printing multiple pages comprising:  
means for selecting a page;  
5 means for automatically determining one or more pages related to said selected page;  
and means for printing said selected page and said one or more related pages.
- 10 2. The apparatus of claim 1, further comprising a page parsing and listing mechanism.
3. The apparatus of claim 1 or 2 wherein the selected page is selected using a Uniform Resource Locator (URL).
- 15 4. The apparatus of claim 3, wherein a page is related to said selected page if they both reside on the same server.
5. The apparatus of claim 4 wherein a page is related to said selected page if they both have the same base address.
- 20 6. The apparatus of any preceding claim wherein the selected page is a hypertext markup language (HTML) page.
- 25 7. The apparatus of any preceding claim, wherein said apparatus comprises a computer program product comprising computer program instructions recorded on a storage medium.
8. The apparatus of any of claims 1 to 6, wherein said apparatus  
30 comprises a computer system including at least one processor, a memory coupled to the at least one processor; and computer program instructions in said memory for execution by the program to perform said printing operation.
- 35 9. An apparatus including a web page print mechanism comprising:  
a web page selection mechanism that allows a user to select at least one web page from the list of web pages;  
a web page parsing and listing mechanism that generates a list of web pages related to the selected web page;  
40 and a mechanism for printing the at least one web page selected by the user using the web page selection mechanism.